AD-754 441

INTERFACE MESSAGE PROCESSORS FOR THE ARPA COMPUTER NETWORK

Frank E. Heart

Bolt Beranek and Newman, Incorporated

Prepared for:

Advanced Research Projects Agency

January 1973

Report No. 2499 January 1973

INTERFACE MESSAGE PROCESSORS FOR THE ARPA COMPUTER NETWORK

QUARTERLY TECHNICAL REPORT NO. 16

1 October 1972 to 31 December 1972

Principal Investigator: Mr. Frank E. Heart
Telephone (617) 491-1850, Ext. 470

Sponsored by
Advanced Research Projects Agency
ARPA Order No. 1260

Contract No. DAHC15-69-C-0179
Effective Date: 2 January 1969
Expiration Date: 31 December 1973
Contract Amount: $7,517,008

Title of Work: IMP

Submitted to:

Director
Advanced Research Projects Agency
Arlington, Virginia 22209

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Mass. 02138	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

QUARTERLY TECHNICAL REPORT NO.]6

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Bolt Beranek and Newman Inc.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1973	~~41~~ 44	

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) BBN Report No. 2499
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency Arlington, Virginia 22209

13. ABSTRACT

The basic function of the ARPA computer network is to allow large existing computers (Hosts), with different system configurations, to communicate with each other. Each Host is connected to an Interface Message Processor (IMP), which transmits messages from its Host(s) to other Hosts and accepts messages for its Host(s) from other Hosts. There is frequently no direct communication circuit between two Hosts that wish to communicate; in these cases intermediate IMPs act as message switchers. The message switching is performed as a store and forward operation. The IMPs regularly exchange information which: allows each IMP to adapt its message routing to the conditions of its local section of the network; reports network performance and malfunctions to a Network Control Center; permits message tracing so that network operation can be studied comprehensively; allows network reconfiguration without reprogramming each IMP. The Terminal IMP (TIP), which consists of an IMP and a Multi-Line Controller (MLC), extends the network concepts by permitting the direct attachment (without an intervening Host) of up to 64 dissimilar terminal devices to the network. The Terminal IMP program provides many aspects of the Host protocols in order to allow effective communication between a terminal user and a Host process. A High Speed Modular IMP (HSMIMP) is under development; one goal of this effort is to increase IMP performance by a factor of 10.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Computers and Communication						
Store and Forward Communication						
ARPA Computer Network						
Interface Message Processor						
IMP						
Terminal IMP						
TIP						
Honeywell DDP-516						
Honeywell 316						
Multi-Line Controller						
MLC						
Network Control Center						
NCC						
Host Protocol						
High Speed Modular IMP						
HSMIMP						
Lockheed SUE						
Network Throughput						

ib

Report No. 2499 Bolt Beranek and Newman Inc.

INTERFACE MESSAGE PROCESSORS FOR THE ARPA COMPUTER NETWORK

QUARTERLY TECHNICAL REPORT NO. 16
1 October 1972 to 31 December 1972

Submitted to:

Advanced Research Projects Agency
Arlington, Virginia 22209
Attn: Dr. L.G. Roberts

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC-15-69-C-0179.

ic

TABLE OF CONTENTS

1. OVERVIEW

This Quarterly Technical Report, Number 16, describes aspects of our work on the ARPA Computer Network during the last quarter of 1972.

During this quarter we installed two 316 IMPs and two TIPs and relocated a 516 IMP. The 316 IMPs were installed at the University of California at San Diego and at the Rand Corporation. The 516 IMP which was previously installed at the Rand Corporation was moved to the Information Science Institute of the University of Southern California. One TIP was installed at the University of Hawaii; this system is connected to the rest of the network via an earth-satellite communication circuit. A second TIP was temporarily installed at the 1972 International Conference on Computer Communication (in Washington, D.C.) for a three-day demonstration of the network. Following the conference, this TIP was installed at Fleet Numerical Weather Central in Monterey California.

During the fourth quarter we completed a study of the theoretical and measured throughput of the ARPA network. This study brings together the results of several investigations carried out during 1972 both by BBN and by other groups. The results of this work are presented in Section 2.

1.1 IMP/TIP Development

The software for both IMPs and TIPs continued to evolve during the fourth quarter. The largest single change to the IMP program was in the mechanisms used to detect dead Hosts. The IMPs now place slightly less reliance on a Host's *Ready line* (which several Hosts were using improperly) and somewhat more reliance on timing messages crossing the Host/IMP interface. In

addition, source Host notification of a dead destination Host has been moved from the source to the destination IMP; this change was made partially for reduction of current IMP table size, and partially in preparation for "area routing".

Most of the TIP software changes were in the nature of further tailoring of the TIP code to meet the desires of particular sites. One new TIP command (RESET terminal parameters) was added. The TIP NEWS feature, which actually resides on a TENEX system, was expanded to provide Host survey information and to accept user comments and complaints. In addition, we began the process of making the *TIP User's Guide* available on-line at BBN (TENEX) and at the Network Information Center.

In our Quarterly Technical Report Number 12 we discussed a tentative approach to the connection of remote batch terminals to the Multi-Line Controller (MLC) of the TIP. We continued to investigate this approach during 1972; this effort culminated in the production of a complete specification in September and the solicitation of proposals from several vendors early in the fourth quarter. After a careful evaluation of the nine proposals we received, it was decided that the costs, both for development and for subsequent purchase of additional units, were incompatible with the perceived benefits to be obtained. Therefore, after obtaining ARPA concurrence, we have abandoned this approach and are now investigating the interposition of a "mini-Host" between a remote batch terminal and an IMP (or TIP). The mini-Host will be built from the same modular equipment being used for construction of the HSMIMP and will fit smoothly into longer-term plans for HSMIMP architecture. The first version of the mini-Host is likely to be tailored for connection of the IBM 2780 (or equivalent), but could later be tailored for other remote batch terminals.

During the fourth quarter we intensified our investigation of modems which might be connected to the TIP's MLC. The TIP currently supports Bell 103 modems (and equivalents); these modems are full-duplex but limited to a maximum rate of 300 baud. There is expanding interest, on the part of TIP users, in modems which can operate at higher speed over the dial network. Two areas of particular interest are simplex modems (for attachment of output-only devices such as line printers), and modems with a high data rate in one direction and a moderate rate in the reverse direction (e.g., 150 baud input, 1200 baud output) for connection of keyboard/display devices. We have begun experiments with a number of such modems to determine what modifications, if any, will be required in the TIP code.

1.2 Network Control Center

A number of changes were made in the Network Control Center (NCC) during the quarter. The single largest change was the addition of a "Host software consultant" to the NCC staff. The notion is to provide a single point which network users can contact to obtain software consulting services. Naturally, no single individual can be intimately familiar with all of the systems and subsystems available through the network; however, we hope that an individual whose primary concern is answering users' questions will be able to provide quicker and more reliable consulting than was previously available. Other NCC changes include the restructuring of the NCC program to take note of network partitions and the continued refinement of operator procedures.

1.3 HSMIMP Development

Work continued on the design and construction of the High Speed Modular IMP (HSMIMP) during the fourth quarter. By the end of the quarter we had received seventeen processors, twenty 4K memories, and five 8K memories plus auxiliary equipment from Lockheed; this is somewhat more than the amount of Lockheed equipment required to fabricate one full-scale HSMIMP. We have designed and fabricated prototypes of the bus coupler, the clock, the Priority Interrupt Device (PID), the DMA, and versions of the Host and modem interfaces compatible with the current IMP systems. Of these, debugging of the clock and PID units has essentially been completed.

With regard to HSMIMP software, the store-and-forward routines and the DDT have been coded and debugging will begin in 1973.

1.4 Publications and Conference Participation

As part of our technical interaction with other groups, a notable activity was our participation in the 1972 International Conference on Computer Communication (ICCC). A Terminal IMP, connected to the network by two 50-kilobit circuits, was installed at the conference and used for both prepared and hands-on demonstrations of the network. About 25 different terminal types were loaned by their manufacturers for this demonstration; many of these had not previously been connected to the TIP. In spite of the relatively short time available for experimentation and debugging, all were operated successfully during the demonstration. In addition to this demonstration, we presented a paper at the ICCC entitled *The Network Control Center for the ARPA Network*.

Other publications during the fourth quarter include presentation of a paper entitled *Improvements in the Design and Performance of the ARPA Network* at the 1972 Fall Joint Computer Conference and the submission of papers to the 1973 Hawaii System Science Conference and to COMPCON 73. Finally, a revision to BBN Report No. 1822, *Specifications for the Interconnection of a Host and an IMP,* was distributed at the end of the quarter.

2. NETWORK THROUGHPUT

This Section brings together the results of several investigations into throughput in the ARPA Network carried out over the course of the past year. The theoretical results and most of the measurements were developed by BBN; some of the measurements have been reported by other groups in the Network. Some of the results can be summarized as follows:

- The throughput of the IMP processor is shown to be an increasing function of message length. For full length messages, the IMP can process one megabit per second of line traffic. It can process almost half a megabit per second of local Host traffic.
- The IMPs can support between 35 and 40 kilobits per second of Host data on a single 50-Kbs circuit or paths of several 50-Kbs circuits. The issues of larger paths and multiple paths are explored.
- For short paths, a Host can attain 35 to 40 Kbs of throughput using one message at a time. For paths with more than about 8 hops, the Host is reduced to 50% throughput with only one message. For paths with satellite links, the Host must use 5 or more messages in flight to maintain 40 Kbs of throughput.
- The buffering needed to drive a terminal at maximum rate is analyzed as a function of distance. A 300-baud terminal requires only 6 characters to buffer it for most current Network paths A 2400-baud terminal requires 50 characters or more for the same path, and a buffer of nearly 250 characters to go over a satellite.
- Several Host measurements are reported, indicating that Hosts have been able to achieve throughput rates of 25 to 35 Kbs over sustained periods, but that this is usually very costly in Host processing time.

Preceding page blank

2.1 Theoretical Considerations

We begin by describing several factors that determine the throughput available to a Host on the ARPA Network. The functional dependence between these parameters and network performance is discussed to provide a background for the experimental results presented later.

2.1.1 IMP Processor Throughput

To calculate the maximum throughput (in Host data bits processed per second) that an IMP can support, we consider the computational load placed on the IMP by the processing of one message. This load must take into account machine instruction cycles and input/output cycles required to process all the packets of a message, their acknowledgements, the RFNM for the message, and its acknowledgement. This analysis was carried through in [1], and we summarize the results in Figure 1. There is an overhead on each packet and an overhead on the message as a whole, so it is reasonable that there are discontinuities at packet boundaries and that full length messages are the most efficient. The higher curves plot actual number of bits per second on the phone line; the lower curves plot Host data bits per second. The difference between the two is a measure of overhead. Notice that a 516 IMP has the processing capability to handle 20 50-Kilobit circuits full duplex (a megabit per second) when operating on full-length messages. The difference between the 516 and 316 IMP processors is in the core memory cycle time and memory channel speed.

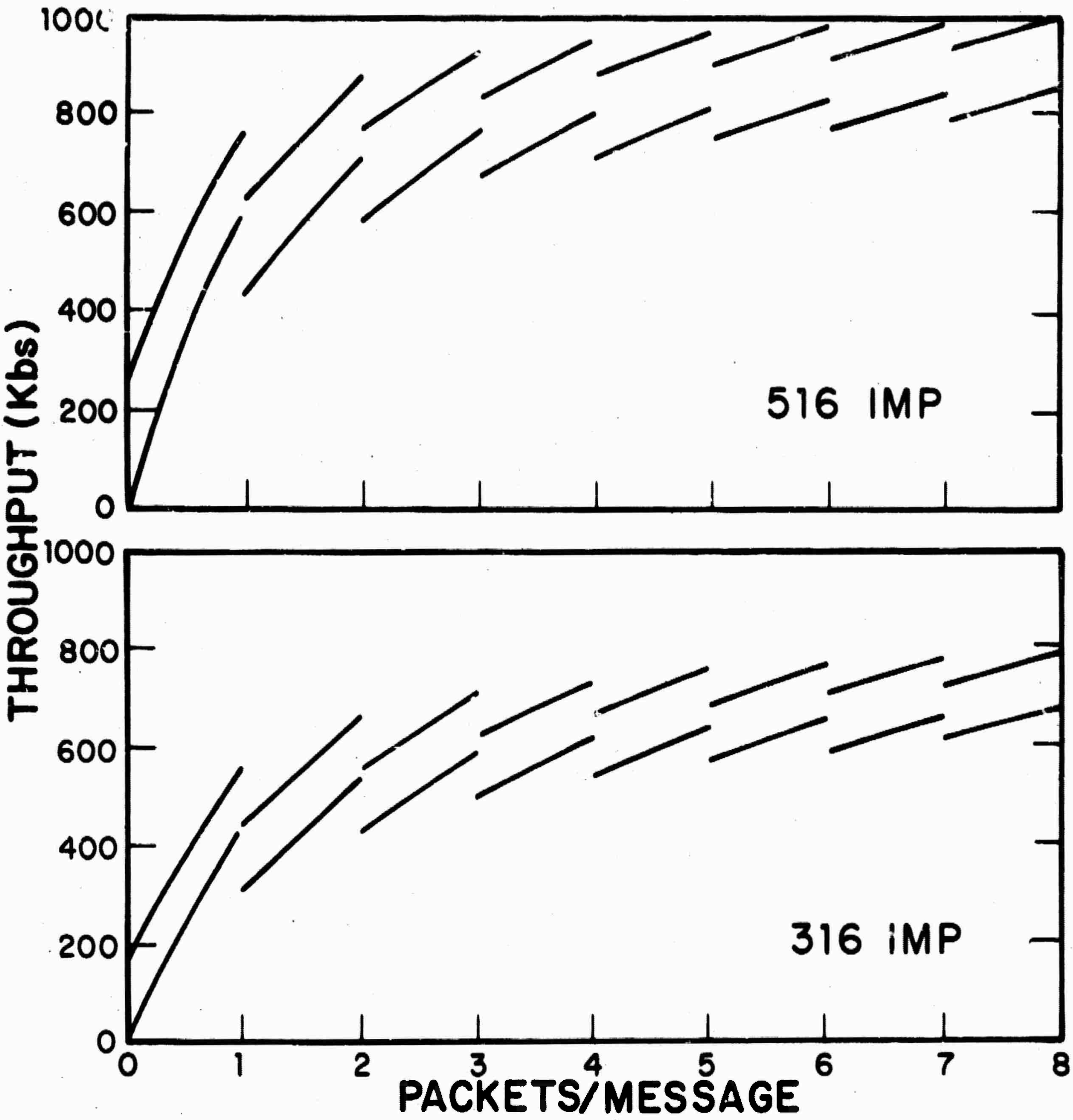

FIGURE 1. IMP PROCESSOR THROUGHPUT VS. MESSAGE LENGTH

The higher curves plot Line Capacity, the lower curves plot Net Throughput

2.1.2 IMP-IMP Throughput

The fraction of useful bandwidth available to the IMP sub-network from a communications circuit can be computed as follows:

1. The IMP sends routing messages every 2/3 of a second, which requires about 2kbs of bandwidth, independent of the line speed.

2. For full length packets (1008 bits), the software and hardware overhead totals to 16%. These two factors reduce the useful bandwidth of a 50-kbs circuit to 40 kbs, and a 230.4-kbs circuit to 191 kbs.

The useful bandwidth obtainable from a circuit also depends on the line length and the error control strategy. The IMP buffers each packet that it transmits until it receives an acknowledgement, meanwhile transmitting other packets to utilize the circuit efficiently. If it does not receive an acknowledgement in the expected time, it retransmits the packet. The expected time for an acknowledgement to return is the sum of:

1. Speed-of-light propagation delay for the packet--the time for the first bit to traverse the circuit, a function of distance.

2. Transmission delay for the packet-- the time for the bits of the packet to be clocked onto the circuit, a function of its bandwidth.

3. IMP processing delay.

4. Latency in sending the acknowledgement--queueing delay plus delay for the transmission currently in progress.

5. Propagation delay for the acknowledgement.

6. Transmission delay for the acknowledgement.

7. Other IMP processing delay.

Thus, for a longer line or a higher speed line, the IMP must buffer proportionately more packets if the maximum bandwidth of the line is to be obtained. The IMP buffers up to 8 packets for all land lines and 32 packets will suffice for satellite links. On these grounds, all IMPs connected to satellite links will have extra core. A final note about satellites--since many IMPs will be competing for the use of a single satellite, the discussion above is overly simplified to be useful in the analysis of the bandwidth obtainable from a satellite. In fact, there are certainly extra costs that we have not detailed here which further reduce the useful bandwidth. A fuller explanation of satellite communication can be found in [2] and [3].

We can compute the number of packet buffers needed to fully utilize a circuit of any speed and length. Figure 2, taken from [1], shows that this number also depends on the traffic mix. It has a linear dependence on line speed and line length, plus a constant term. The knee of the curve occurs at shorter distances with higher speeds, and the constant term is insignificant at very high speeds. (A short packet contains 152 bits of overhead and no data, a long packet carries an additional 1008 bits of data for a total of 1160 bits.)

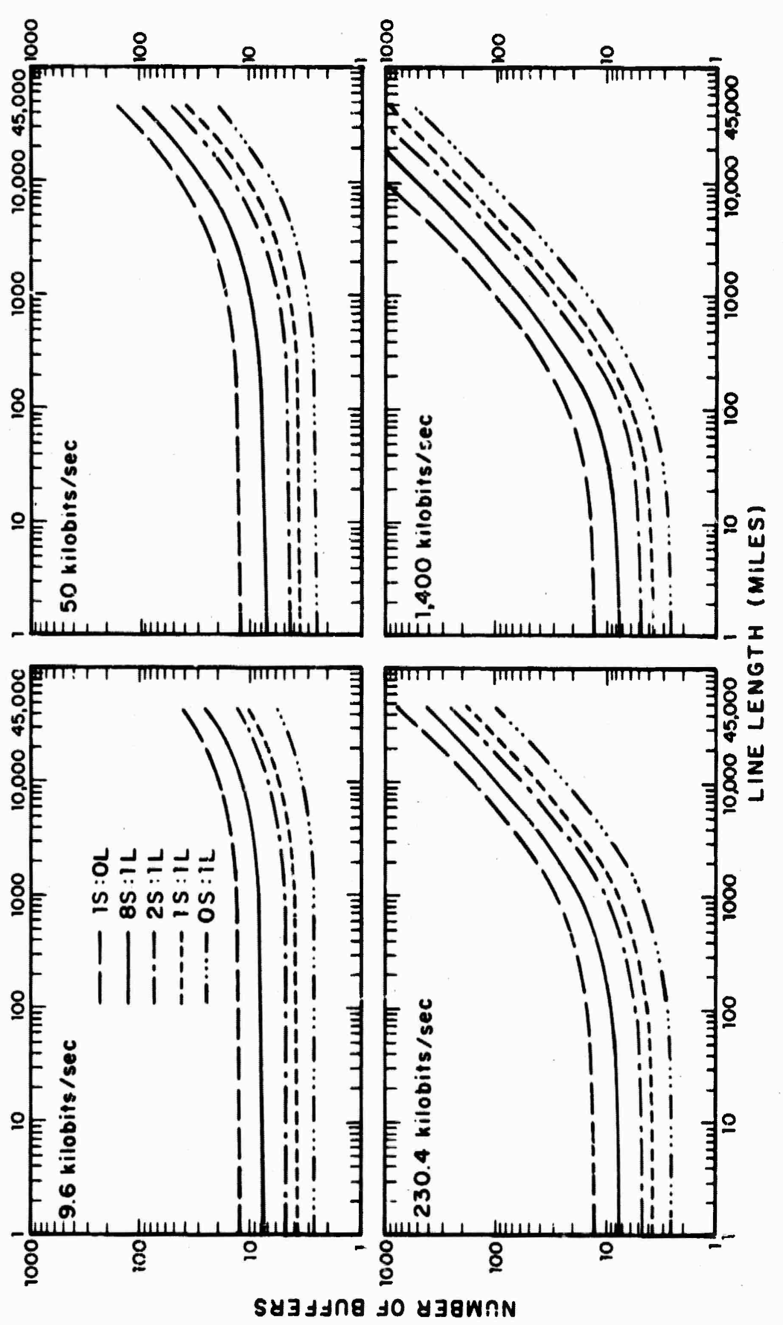

FIGURE 2. NUMBER OF BUFFERS FOR FULL LINE UTILIZATION
Traffic mixes are shown as the ratio of numbers of short packets (S - zero data bits) to number of long packets (L - 1008 data bits).

2.1.3 Host-Host Throughput

We now consider the question of the throughput that a pair of Hosts in the ARPANET can sustain between themselves. We postpone discussion of the effects that are due to multiple network paths between the two Hosts. For this discussion, we take the maximum throughput attainable by a pair of IMPs to have the value 1, and we consider what fraction of the maximum throughput is available to a pair of Hosts. To sustain the maximum rate, the source Host must keep the virtual end-to-end path filled with data at all times. We introduce the number

$$F = \frac{\text{round trip time}}{\text{single message time}}$$

which is the number of messages needed to fill the entire path and therefore reach maximum throughput. Notice the close analogy between the number of messages needed to buffer a source-to-destination network path (considered here) and the number of packets needed to buffer a single IMP-to-IMP circuit (considered in the previous section). We will consider only full-length messages because they utilize the resources of the network most efficiently.

We assume that IMP processing delays are small for each packet and ignore the effects of RFNMs, we have

$$F = \frac{H x (T+[2xL])+ 7xT}{8xT}$$

where H = the number of hops in the end-to-end path
T = the transmission time of one packet over one hop
L = the speed of light propagation delay per hop

Table 1 shows the dependence of F on the line characteristics and the number of hops.

TABLE 1

NUMBER OF FULL-LENGTH MESSAGES NEEDED TO ACHIEVE MAXIMUM THROUGHPUT

LINE TYPE	NUMBER OF HOPS 1	2	4	8	16	32
NO SATELLITE LINKS						
10MI 9.6 kbs	1.0	1.1	1.4	1.9	2.9	4.9
10MI 50.0 kbs	1.0	1.1	1.4	1.9	2.9	4.9
10MI 230.4 kbs	1.0	1.1	1.4	1.9	2.9	5.0
10MI 1400.0 kbs	1.0	1.2	1.4	2.0	3.1	5.4
100MI 9.6 kbs	1.0	1.1	1.4	1.9	2.9	4.9
100MI 50.0 kbs	1.0	1.1	1.4	1.9	3.0	5.1
100MI 230.4 kbs	1.0	1.2	1.5	2.1	3.3	5.7
100MI 1400.0 kbs	1.2	1.5	2.0	3.2	5.5	10.1
1000MI 9.6 kbs	1.0	1.1	1.4	2.0	3.1	5.2
1000MI 50.0 kbs	1.1	1.2	1.6	2.3	3.8	6.7
1000MI 230.4 kbs	1.3	1.7	2.4	4.0	7.2	13.5
1000MI 1400.0 kbs*	2.6	4.4	7.9	14.9	28.9	57.0
3000MI 9.6 kbs	1.0	1.2	1.5	2.1	3.4	5.9
3000MI 50.0 kbs	1.2	1.5	2.1	3.3	5.7	10.5
3000MI 230.4 kbs	1.8	2.7	4.6	8.3	15.7	30.6
3000MI 1400.0 kbs*	5.9	10.9	20.9	41.0	81.1	161.3
ONE SATELLITE LINK (SAME RATE AS THE LAND LINES)						
10MI 9.6 kbs	2.5	2.6	2.9	3.4	4.4	6.4
10MI 50.0 kbs	4.6	4.7	5.0	5.5	6.5	8.5
10MI 230.4 kbs	14.1	14.2	14.5	15.0	16.0	18.0
10MI 1400.0 kbs	75.3	75.5	75.8	76.3	77.5	79.7
100MI 9.6 kbs	2.5	2.6	2.9	3.4	4.4	6.4
100MI 50.0 kbs	4.6	4.8	5.0	5.5	6.6	8.7
100MI 230.4 kbs	14.1	14.2	14.5	15.2	16.4	18.8
100MI 1400.0 kbs	75.5	75.8	76.3	77.5	79.8	84.4
1000MI 9.6 kbs	2.5	2.7	2.9	3.5	4.6	6.7
1000MI 50.0 kbs	4.7	4.9	5.2	6.0	7.4	10.4
1000MI 230.4 kbs	14.3	14.7	15.5	17.1	20.2	26.5
1000MI 1400.0 kbs*	76.9	78.7	82.2	89.2	103.3	131.3
3000MI 9.6 kbs	2.5	2.7	3.0	3.6	4.9	7.5
3000MI 50.0 kbs	4.8	5.1	5.7	6.9	9.3	14.1
3000MI 230.4 kbs	14.9	15.8	17.7	21.4	28.8	43.7
3000MI 1400.0 kbs*	80.2	85.2	95.2	115.3	155.4	235.6

*(Land circuits at megabit rates are not currently available over long distances.)

We can consider some of the consequences of the results shown in Table 1. Suppose that two Hosts are separated by a network path of 8 lines with each 50-kbs circuit about 100 miles long. In this case, two messages are required to obtain the theoretical maximum throughput. If the Hosts are following the standard Host-Host protocol, they will find a reduction to 50% of the maximum throughput due to the one message per link rule [7]. Of course, by using several links, a Host can overcome this problem. *Note that the throughput rates obtained by a pair of Hosts may be below maximum while the IMP-IMP throughput along the path is maximum if there is sufficient store-and-forward buffering.* The important fact illustrated in Table 1 is that the number of messages needed to buffer most current network paths that do not include satellite links is less than 3. However, the introduction of a satellite link at 50-kbs immediately adds 3 to this number. This means that the Host must be able to sustain as many as 5 or 6 full-length messages in flight all the time. Of course, this number also goes up with other measures of distance, like line speed and the number of hops.

There also are implications for the IMP subnetwork in these figures. If the Host must send off several messages to achieve high throughput, then the IMP must buffer these messages and do various kinds of bookkeeping for them. A more complete discussion of some of these issues can be found in Section 2.3.

It is clear that the strategies used by the IMP and Host systems for achieving good throughput should take account of the network topology. The network has grown a great deal in the last year or two, but a much more fundamental change will take place with the introduction of satellite links and higher speed circuits.

2.1.4 Host-Terminal Throughput

Another type of throughput that is interesting is the throughput between a terminal on one Host (for instance on a TIP) and another Host across the network. The main issue here seems to be the minimum size of the terminal buffers which will permit the terminal to operate at its maximum rate. In the rest of this section we compute this minimum buffer size for a number of terminal speeds and for a variety of "distances" between the terminal and the Host. We consider 12 terminal rates in the range 75-baud to 19200-baud. We will use the hypothetical network path shown in Figure 3.

The Host is at the left side of the figure and the terminal may be connected to a Host on any of the IMPs except the left-most IMP. In other words, Figure 3 represents a variety of typical distances between a Host and a terminal across the network. For instance, the first three or four hops represent "across town" in Boston, Washington, Palo Alto, or Los Angeles. The first five hops represent across town and up or down the coast (Los Angeles to Palo Alto, Boston to Washington). The first six hops represent across town, up or down the coast, and to the midwest. The first seven hops represent across town, up the coast, and across country. Adding the eighth hop represents adding a satellite hop to Hawaii or Europe, and the ninth hop represents getting from the satellite ground station to a Host. The tenth hop represents the addition of a cross Europe link (for example Oslo to London). For simplicity, all circuits are assumed to be 50-kbs. If the satellite and the following two hops run at 9.6-kbs (as they may in Europe), there is no change below 300-baud. From 300-baud to 2400-baud, the buffer requirements are increased by approximately 50%, and terminals with higher speeds cannot be supported at their maximum rates.

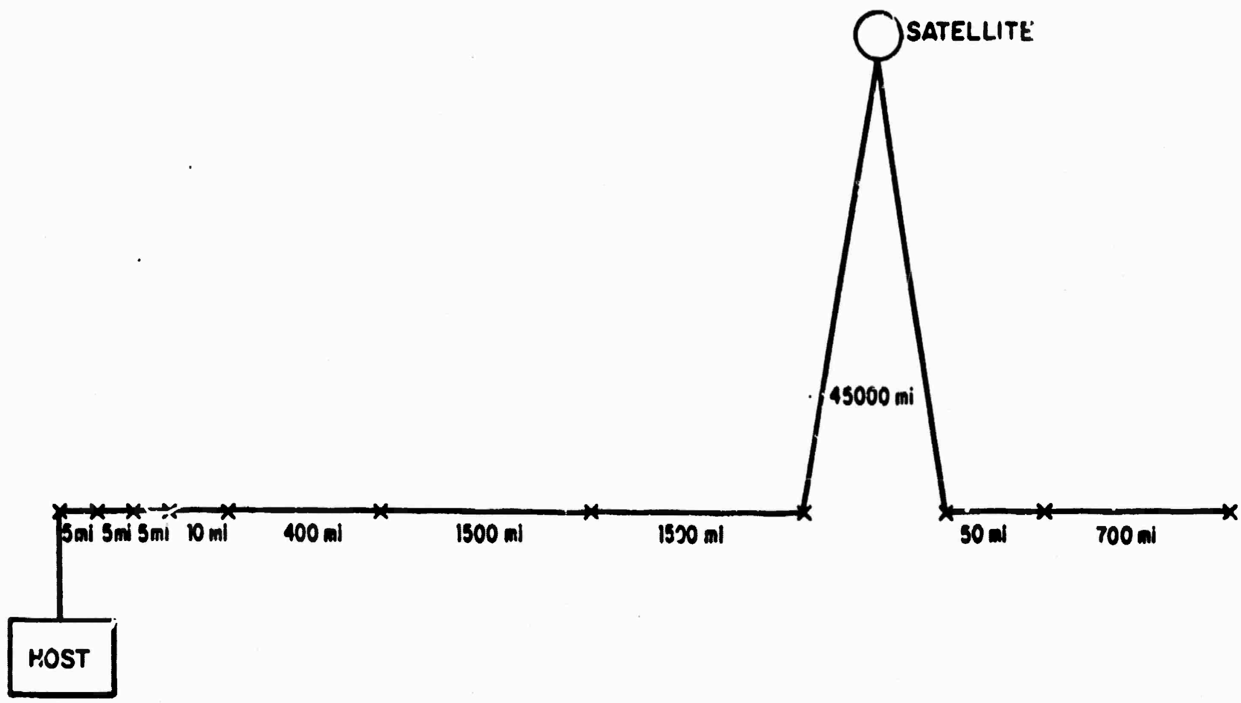

FIGURE 3. A HYPOTHETICAL NETWORK PATH

Each "x" represents an IMP

The minimum buffering required to keep a terminal printing at top speed is a function of terminal rate and network delay. For some optimum length message, the network round trip time is equal to the time taken by the terminal to print the characters of the message. We calculate the network round trip time as a function of message length in a manner similar to [1] except that we use the transmission delay for a Host/Host protocol allocate instead of the RFNM. Also, we assume 8 bit characters, and a Host processing time of 100 milliseconds per message before the allocate is returned. Message sizes between 2 and 990 characters are allowed, and we assume that a single Host/Host protocol allocate is sufficient for each message. For the other half of the calculation, we take the printing time for a message of C characters at a baud rate of R to be

$$\frac{10 \mathrm{x} C}{R}$$

That is, we take into account start and stop bits and say that the 8 data bits of the character turn into 10 bits when sent to the terminal.

Table 2 shows the results obtained by solving the equation between network round trip delay and message printing time for minimum buffer size. While the entries in the table are the minimum size buffers to maintain full terminal throughput, a standard double buffer implementation such as is used in the TIP requires twice the buffering given in the table. It states, for instance, that terminals at rates of 300-baud or less need 6 or less characters of buffering for most terrestial network paths. Even at 2400-baud, a character buffer of about 50 will usually suffice, although at that terminal speed, the buffering is indeed a function of the length of the network path. Most noticeably, however, the buffering needs go up as a satellite link is introduced. To keep a 2400-baud terminal printing at top speed over a satellite link requires roughly 3 to 5

times the buffering needed otherwise--nearly 250 characters. (It should be noted that these storage requirements for terminals are often not the limiting factors on the *number of terminals* which can be connected to a Host or a TIP. There are often hard constraints on the numbers of active jobs or connections on such a system and, certainly, processing capability limitations. For instance, it is reported in [4] that it takes 15% CPU utilization on TENEX to keep a 2400-baud terminal printing at top speed.)

TABLE 2

SIZE OF THE CHARACTER BUFFER NEEDED TO DRIVE A TERMINAL AT ITS MAXIMUM RATE

TERMINAL RATE (baud)	MILEAGE INCREMENT FOR EACH HOP (length of the network path is cumulative from left to right)									
	5	5	5	10	400	1500	1500	45000	50	700
75.0	<2	<2	<2	<2	<2	<2	<2	6	6	6
110.0	<2	<2	<2	<2	3	3	3	10	10	10
134.5	<2	<2	3	3	3	3	3	10	10	10
150.0	<2	3	3	3	3	3	3	10	10	14
300.0	3	6	6	6	6	6	6	22	26	26
600.0	6	10	10	10	10	14	18	48	52	52
1200.0	14	18	18	22	26	30	36	106	114	118
1800.0	22	26	30	33	40	48	56	164	172	180
2400.0	30	36	40	48	56	70	82	222	230	246
4800.0	60	74	94	125	144	168	194	466	494	516
9600.0	152	184	218	276	314	384	432	>990	>990	>990
19200.0	404	520	634	750	870	>990	>990	>990	>990	>990

2.2 Measurements

This section contains some recent measurements of throughput in the ARAPNET. A brief description is given of the methods employed, then a summary of the results is provided, together with some conclusions and some rules of thumb.

2.2.1 IMP Subnetwork Measurements

To calculate the throughput attainable in the IMP subnetwork, we make use of the Fake Host programs in the IMP. Specifically, we use Message Generator to send artificial traffic to Discard. These are programs which simulate the action of the hardware in transferring real Host messages in and out of memory. They run on a word-by-word basis: Message Generator takes 25 cycles per 16 bit word; Discard takes 20 cycles. These programs utilize all the IMP program mechanisms for processing messages and are therefore comparable in every way with real Host traffic (with two exceptions--they are very fast, and they are more taxing of IMP processor bandwidth than hardware Host transfers).

All the IMP subnetwork measurements cited below were obtained by sending 256 full length messages (just over 2 megabits) from Message Generator in one IMP to Discard in some other IMP and measuring the elapsed time. Finally, these measurements were made in the network while actual traffic was flowing from Host to Host. One must not neglect to account for the steady-state use of all lines in the network. Current measurements indicate that lines are utilized between 1% and 10% on a long-term average, with the average use about 3.5%. We expect this rate to increase significantly as network use expands.

2.2.1.1 Internal Throughput

There are two special cases which we will discuss first because they isolate certain phases of IMP processing and provide a means of comparison between the different IMP processors. These cases are messages to Hosts at dead or nonexistent IMPs and messages to Hosts at the same IMP as the source. In the first case, only the Host input routine is run. In the second case, only Host input, Host output and TASK are run. Thus, these cases serve to calibrate the processing speed of the IMP. The results of some measurements on the different IMP processors are presented in Table 3A.

The discrepancy between the performance of the 516 IMP and the 316 IMP is explained exactly by the difference in cycle time of the memory, 1 μsec for the 516 and 1.6 μsec for the 316. The TIP, however, is 35% slower than the 316 IMP, giving an effective cycle time of 2.2 μsec. This figure represents the extra processing time required by a relatively inactive TIP. This percentage more than doubles under heavy load. (It should be noted that certain critical portions of the TIP program are currently being redesigned and these measurements may be significantly improved.)

Since we know exactly how much processing time the Message Generator and Discard processes require, we can determine how much processing time is used for actual message handling in the IMP and how much time is taken up by the hardware-simulating processes. The results show that, in the dead case, only 20% of the processing is in the IMP itself, and in the local case, only 45%. These calculations are summarized in Table 3B.

TABLE 3

INTERNAL THROUGHPUT RATES

		Source		
		516 IMP	316 IMP	316 TIP
Destination	Dead	505	295	230
	Self	195	120	90

A. Measured Internal Throughput Rates (in kbs)

		Source		
		516 IMP	316 IMP	316 TIP
Destination	Dead	2525 (631)	1475 (368)	1150 (287)
	Self	433 (354)	266 (218)	200 (165)

B. Derived Internal Throughput Rates (in kbs)
IMP Processing alone and (Message Generator/Discard alone)

2.2.1.2 Throughput Over a Single Circuit

Now we will consider the case of messages sent to an IMP one or more hops away in the network. We first address the single hop case and investigate how closely the IMPs come to the 40 kbs maximum rate for 50 kbs circuits. It is interesting to compare the different IMP processors to show that the great variance in processing power that was evident in the internal case is not a dominant factor in the case of 50 kbs circuits. The measurements were made in two ways: one-way traffic over the line and two-way traffic. The results in Table 4 show that the different processors are about the same (the differences between them are small and can almost be accounted for by the level of other traffic).

These figures do not reach the theoretical limit of 40 kbs because of the steady-state level of other traffic in the network. Measurements with IMPs on spur connections confirm this observation, since they usually attain 40 kbs.

To a first-order approximation, it makes no difference whether it is an IMP or a TIP, a 516 or a 316, at the source or the destination. Two-way traffic at capacity levels tends to reduce throughput rates by less than 10% due to RFNM processing and reverse direction message processing delays.

TABLE 4: THROUGHPUT RATES OVER 50 KILOBIT CIRCUITS

One-Way Traffic (Two-Way) over a Single 50 kbs Line

		Source		
		516 IMP	316 IMP	316 TIP
	516 IMP	38 (36)	38 (35)	38 (34)
Destination	316 IMP	37 (34)	36 (35)	37 (34)
	316 TIP	35 (34)	37 (34)	37 (34)

2.2.1.3 Throughput Over Several Hops

For more than one hop between IMPs, the results are much more complicated. All measurements seem to have a large variance, probably because of the fact that the network is never idle. They seem to indicate:

(1) The same high throughput levels for one hop are sustained on paths which are 4 to 8 hops long.

(2) At some length, the throughput begins to drop off to the range 30-35 kbs. This usually happens from 5 to 10 hops away.

(3) At greater than 10 hops, throughput varies greatly, from 20-35 kbs.

(4) The position of the source IMP in the network is a first-order determinant of throughput performance. IMPs on spur connection perform best (since that single line is dedicated to the experimental traffic) and relatively idle IMPs do somewhat better than busy IMPs.

(5) Some lines are used much more than others, and maximum throughput rates can never be achieved on these lines because of the presence of other traffic.

(6) The degradation of throughput rates with distance is probably related to the observations in (4) and (5) rather than any intrinsic network parameters. That is, the more hops the traffic traverses, the more interference it encounters, and the maximum throughput rate declines.

The distribution of path lengths to various destinations around the ARPA Network changes from IMP to IMP and also as the topology changes. Table 5A shows a typical example, the path

lengths from the BBN IMP to each of the other IMPs in the network on January 1, 1973 with all lines in the network up. For comparison, Table 5B shows the same distribution with the BBN-Harvard line down. Notice that about two-thirds of the IMPs are 6 hops or less from BBN in the first case, while only one-third of the IMPs are within 6 hops in the second case. Of course, the path lengths might increase further in the event of other line failures, the failure of an IMP, or multiple IMP and line failures.

TABLE 5

NETWORK PATH LENGTHS FROM THE BBN IMP*

NUMBER OF HOPS AWAY	NUMBER OF IMPS	FRACTION OF TOTAL	CUMULATIVE NUMBER OF IMPS	CUMULATIVE FRACTION OF TOTAL
1	3	9%	3	9%
2	2	6%	5	15%
3	3	9%	8	24%
4	3	9%	11	33%
5	4	12%	15	44%
6	6	18%	21	64%
7	8	24%	29	88%
8	3	9%	32	97%
9	1	3%	33	100%

A. Network Path Lengths from the BBN IMP--All Lines Up

NUMBER OF HOPS AWAY	NUMBER OF IMPS	FRACTION OF TOTAL	CUMULATIVE NUMBER OF IMPS	CUMULATIVE FRACTION OF TOTAL
1	2	6%	2	6%
2	1	3%	3	9%
3	2	6%	5	15%
4	2	6%	7	21%
5	2	6%	9	27%
6	4	12%	13	39%
7	5	15%	18	55%
8	5	15%	23	70%
9	6	18%	29	88%
10	4	12%	33	100%

B. Network Path Lengths from the BBN IMP--One Line Down

*January 1, 1973

2.2.1.4 Throughput with Multiple Traffic Sources

In the previous sections, we have considered a single traffic stream on a single circuit and on a series of circuits. We have not addressed the more complex question of several traffic patterns over the same set of lines and IMPs.

We can examine the issue of interference between traffic under controlled circumstances. Measurements were made in the specific situation of multiple IMPs sending artificial traffic to Discard at the same destination IMP. First, we investigated the performance of the network when separate paths exist from each source to each destination. The current network topology permits experiments with 2,3, and 4 IMPs adjacent to a 516 IMP. When a single source IMP was active, the throughput rates averaged between 36-38 kbs. When 2 source IMPs were active at the same time, no degradation was observed. With 3 source IMPs the average throughput fell to 32-33 kbs, and with 4 IMPs the rate dropped to 26-27 kbs. The reason for this reduced performance is clearly the limit on reassembly storage. Referring to Table 1, we see that with only 3 message buffers, the IMP cannot possibly support 4 different traffic streams running at maximum rates on 50 kbs circuits. This conclusion was partially verified by performing the experiment of 3 source IMPs with the BBN IMP as the destination (at this writing, it is the only 516 IMP with 16K of core and 10 message reassembly buffers). With 3 source IMPs, no reduction in throughput was observed.

We next investigated the case of multiple sources sending to the same destination over a shared path. We picked three IMPs connected by two lines and sent traffic from two IMPs to the third. Here one line is used by one IMP and the other is shared by both IMPs. The measured performance revealed that the

the fundamental network algorithms for routing and storage allocation are very fair indeed, because both source IMPs achieved the same throughput level, 20 kbs, which is half of the available bandwidth on the shared circuit. The experiment was repeated with a chain of three source IMPs, and also with two separate chains of two and three IMPs. In each case, the results were similar; all source IMPs received approximately the same fraction of the shared resource, the line bandwidth.

2.2.1.5 Throughput over Multiple Paths

Under some circumstances, the IMP is able to divert an input traffic stream over more than one output line and by this parallelism achieve a higner throughput than is possible with a single line. In the Host-to-Host situation, this means that there have to be two or more completely independent paths from the source to the destination, or else the traffic from end to end is clipped to the maximum level attainable on the common line. In the past, the IMP performed this load splitting in a simple-minded but effective manner. It built up a very long queue for output on one line, then switched to building a queue for another line. The period of this switching was that of the routing computation, 2/3 of a second, since the IMP used only one line at a time for output to a particular destination. The queues were allowed to grow to 30 packets or more in length in order to be able to support full throughput over two lines. There were not enough buffers to make this kind of load splitting work over more than 2 lines. The current IMP program limits its queue lengths to 8 packets and thus can achieve only about 25% extra throughput by using a second line in this manner. In fact, small increases in throughput were observed in the experiments described earlier with several IMPs sending to the same destination. These increases approximated the 25% improvement expected due to load splitting.

All of these results will change in the first quarter of 1973, when we will install a new routing algorithm. This algorithm will explicitly choose from one of several possible lines for output, and thus load splitting will be possible without any long queue buildup. This approach is certainly more efficient, both in terms of IMP buffering and message delay, than the earlier method.

Note that the analysis we carried through in Section 2.1.3 on the number of messages needed for maximum throughput on one path applies to each of the independent paths. To keep three separate paths fully loaded takes the *sum* of the messages required to be in flight on each path.

2.2.1.6 High-Speed Circuits and Satellite Circuits

Finally, there is the question of line speeds other than 50-kbs, and the very much longer network paths associated with the introduction of satellite links into the network. In the realm of different line speeds we cannot do very many experiments, because at the time of this writing there is only one 230.4-kbs circuit in the network (between the Ames IMP and the Ames TIP) and no 9.6-kbs circuits. It is quite difficult to load the fast line with a great deal of actual traffic, since there is only one other circuit into each Ames machine and there is only a single 100-kbs Host at each site. Further, the Message Generator and Discard processes in the TIP have been measured above to run at 120-kbs, far below the 190-kbs theoretical maximum of the line. However, we changed these programs to run a packet at a time rather than a word at a time, and the IMP and TIP were both able to achieve the maximum throughput, transmitting separately and at the same time. More experiments with higher speed lines as they are introduced into the network will be necessary to gauge their effectiveness.

As for the question of satellite links, it is premature to measure the throughput available over the first satellite connection. The implementation plan calls for an expanded number of logical channels for the satellite link (32 rather than 8), as well as a core memory expansion. Then software for broadcast communication will be added, and changes to routing and other algorithms in the network will take place. Only after these changes are complete will it be meaningful to evaluate the performance of the satellite link.

2.2.2 Host Measurements

We now present some Host-Host throughput measurements which have been reported to us. There is no automatic way for us to perform throughput experiments with real Hosts. These results are included to give an indication of some typical results obtained by Hosts in high throughput applications.

2.2.2.1 TENEX Measurements

Here we present some measurements made on user programs running with standard Host-Host protocol. By analogy with our analysis of IMP processing bandwidth in Section 2.1.1 and internal throughput measurements in Section 2.2.1.1, we will begin by presenting some results which calibrate the processing capability of the TENEX system. These results are taken from [4], and the measurements shown in Table 6A were found by a similar procedure to that used for subnetwork measurements. A million bits were sent from TENEX to the IMP and back (using BIN/BOUT and the buffered send mode). There was no other load on the system while the experiment was run.

We measured the TENEX Host interface to be serviced (a word at a time) at a rate of about 75 kbs in both directions for the duration of a message, averaged over 1000 bit-transfers. Thus the difference between this maximum rate and the observed rates is due to inter-mesage processing. Note also that this is strictly a Host measurement and that TENEX performs in a highly responsive manner.

We were not able to perform a set of throughput measurements using several TENEX Hosts on the Network for a variety of reasons, primarily divergent software. Some results were

reported to us [5], however, and they are reproduced here. A byte size of 36 bits was used in all experiments (for 8-bit bytes, the throughput is considerably less). NETSPD is a measurement program which can be directed to use several links while still operating under the basic Host-Host protocol mechanisms on TENEX. FTP is a File Transfer Protocol program which does not use multiple links. Table 6B summarizes these measurements.

TABLE 6

TENEX THROUGHPUT MEASUREMENTS

Data No. of bytes/ byte size	Throughput	CPU Utilization	Derived CPU Bandwidth
50/36	48.8 kbs	16%	291 kbs
100/36	46.5 kbs	15%	309 kbs
200/36	52.7 kbs	16%	322 kbs
400/36	59.3 kbs	22%	267 kbs
800/36	61.1 kbs	21%	279 kbs
50/8	22.0 kbs	31%	70 kbs
100/8	26.7 kbs	35%	76 kbs
200/8	29.0 kbs	36%	79 kbs
400/8	30.0 kbs	37%	81 kbs
800/8	31.9 kbs	40%	78 kbs
1600/8	38.1 kbs	59%	64 kbs
3200/8	41.8 kbs	70%	59 kbs

A. Measurements of TENEX CPU Bandwidth in Local Network Communication

Source Program	BBN-TENEX (0 hops)	SRI (5 hops)	USC (8 hops)
NETSPD--1 link	30-35 kbs	12-16 kbs	10-12 kbs
NETSPD--4 links	50-55 kbs	25-35 kbs	20-25 kbs
FTP--1 link to NIL	20 kbs	N/A	10 kbs
FTP--1 link to Disc	10 kbs	N/A	N/A

B. Measurements of TENEX Throughput over the Network

2.2.2.2 Magnetic Tape Transfer Measurements

There are two different groups which have used the Network to transfer magnetic tapes. Tinker and McClellan (while on the Network) transferred tapes between their Univac 418 Host computer systems. The other application has been the TIPs with the magnetic tape option at ETAC and GWC which have been passing meteorological data back and forth and to CCA TENEX.

The Tinker-McClellan experiments have been reported by the Air Force Communications Service in [6]. These experiments took place over a two month period earlier in 1972, and show an average throughput rate of 25 kbs. Table 7 shows how the throughput varied with the block size used. According to [6], the dips in performance at block sizes of 8.1 kilobits and 15.3 kilobits were due to non-network causes and were expected. The dip at 13.5 kilobits per block is unexplained.

The TIP has not undergone the same kind of systematic test as was performed by the AFCS, but some results are known. The magnetic tape transfers between TIPs average about 10 kbs, using a private protocol based on the Host/Host protocol which allows the use of a single link only. From the TIP to TENEX, it was reported in [4] that a four hour run averaged 7 kbs. It is noted that this accounted for 30% CPU utilization on an unloaded system, which means an ultimate TENEX CPU bandwidth of about 20 kbs for this application.

TABLE 7

THROUGHPUT MEASURED BY THE AIR FORCE COMMUNICATIONS SERVICE

Block size (in Kilobits)	Throughput (in Kilobits/sec)
.9	5.4
1.8	10.0
2.7	14.9
3.6	18.7
4.5	22.2
5.4	24.4
6.3	26.9
7.2	28.8
8.1	22.1
9.0	24.0
9.9	26.1
10.8	26.8
11.7	28.0
12.6	29.0
13.5	28.4
14.4	29.9
15.3	24.4
16.2	26.0
17.1	27.0
18.0	27.5

2.3 Discussion

The results presented in this report affect the design of the IMP and Host communications systems, both in the long term and in the short term. As an example of a long term design decision, we can consider the impact of 1.4 mbs circuits on the IMP. As discussed in Sections 2.1.1 and 2.1.2, the current IMP has neither the processing capability to keep up with such circuits nor the memory to adequately buffer them. The decision was made to construct a new IMP processor (the High Speed Modular IMP) with considerably more processing power and memory, specifically, to be able to support megabit circuits.

It may be obvious that such a dramatic change to the speed of circuits in the Network requires a new design effort in the IMP subnetwork. What may not be so clear is that the current steady growth and evolution of the Network affects many short term design decisions also. Here we may point to the rapid increase in the number of circuits in the network, culminating recently with the satellite link to Hawaii. The length of Host-to-Host paths in the Network has been growing steadily, and with a satellite link, the length of an individual hop may now be increased by more than an order of magnitude.

As we saw in Section 2.1.2, a satellite link requires considerably more buffering to keep it running at full efficiency. Further, we plan to connect several sites to the Network by a single satellite link, and these sites will employ a broadcast communications system to share the available bandwidth of the channel. For these reasons, we plan to give IMPs connected to satellites more memory.

A further development along this line is related to the analysis in Section 2.1.3 of the number of messages needed to

achieve full throughput over the Network. These figures have an impact on several network design parameters. For the IMP sub-network, there are two main considerations. The source-to-destination flow control algorithms described in [1] reserve storage at the destruction IMP before messages are allowed into the network. Currently, the 12K IMP has enough reassembly storage for only 3 full length messages. This is enough messages to buffer most present end-to-end network paths. However, as soon as satellite links (or high speed lines) are introduced, *any potential destination IMP* must have additional core, and any source-destination pair separated by the new line must have a larger message number window, if *full* bandwidth is to be attained. For this reason, all IMPs in the Network are being given an additional 4K of memory, so the IMPs will be able to reassemble as many as 1ι full length messages.

The second IMP constraint that must change is in software rather than hardware. The source-to-destination sequence control assigns a message number to each message, and currently there is an allowed "window" of 4 message numbers which can be active at any time. This window will probably grow at the time that more core is installed.

The growth of the number of messages needed for full throughput also has implications for the Host community. This number can be interpreted as the number of messages that must be in transit between a pair of Hosts to achieve maximum throughput. The present restriction in Host-Host protocol that only one message be outstanding on a given link serves to reduce the throughput that a pair of Hosts obtain over a long network path. In fact, Table 1 shows that this restriction may cut throughput by 50% or more for current network topology. This reduction will be even more dramatic with the introduction of higher speed

lines and satellite links. It is clear that a review of Host-Host protocol issues is necessary to deal with these new developments. This is further indicated by the measurements reported on the processor throughput of IMPs and Hosts. It is clear that the current protocols and techniques for data transfer are very taxing of Host processing and that there is room for improvement in this area. Currently, the Hosts run up against processing limits long before the IMP subnetwork does.

Another area in which IMP subnetwork considerations reach back to affect the Hosts is that of terminal buffering, discussed in Section 2.1.4. Several things should be noted about the minimum buffer sizes calculated above: (1) the Host transmitting to the terminal needs as big a buffer as the terminal requires; (2) these are minimum buffering requirements--if there is any extra delay in the network or the Host processing time is greater than 100 milliseconds, the buffering required is bigger; (3) for the terminal to input at top speed and send to the Host requires the same amount of minimum buffering with the same qualification that if there is any additional delay in the network, additional buffering is required.

In conclusion, the design of the communication systems in the ARPA Network should evolve as the Network grows. Both the IMP and Host systems must adapt to changes in order to maintain the maximum throughput from the Network.

REFERENCES

1. J. McQuillan et al, Improvements in the Design and Performance of the ARPA Network, Proceedings of the AFIPS 1972 Fall Joint Computer Conference.

2. W. Crowther et al, A System for Broadcast Communication: Reservation-Aloha, to be presented at Sixth International Conference on System Sciences.

3. ARPA Satellite System notes 1-29. A privately calculated series of notes contributed to by L. Roberts (ARPA), N. Abramson (U. of Hawaii), L. Kleinrock (UCLA), R. Metcalfe (Xerox PARC), R. Rettberg (BBN), and their colleagues.

4. H. Murray, TENEX Bandwidth, 29 November 1972, RFC 415, NIC 10428. Available from the Network Information Center (NIC).

5. R. Tomlinson and B. Clements, BBN, Personal Communication, October 1972.

6. AFCS Test of the ARPA Network, 20 September 1972, Headquarters Air Force Communications Service, Richards-Gebaur AFB, Missouri.

7. Host/Host Protocol for the ARPA Network, Alex McKenzie, BBN, January 1972, NIC 8246. Available from the NIC.

Preceding page blank